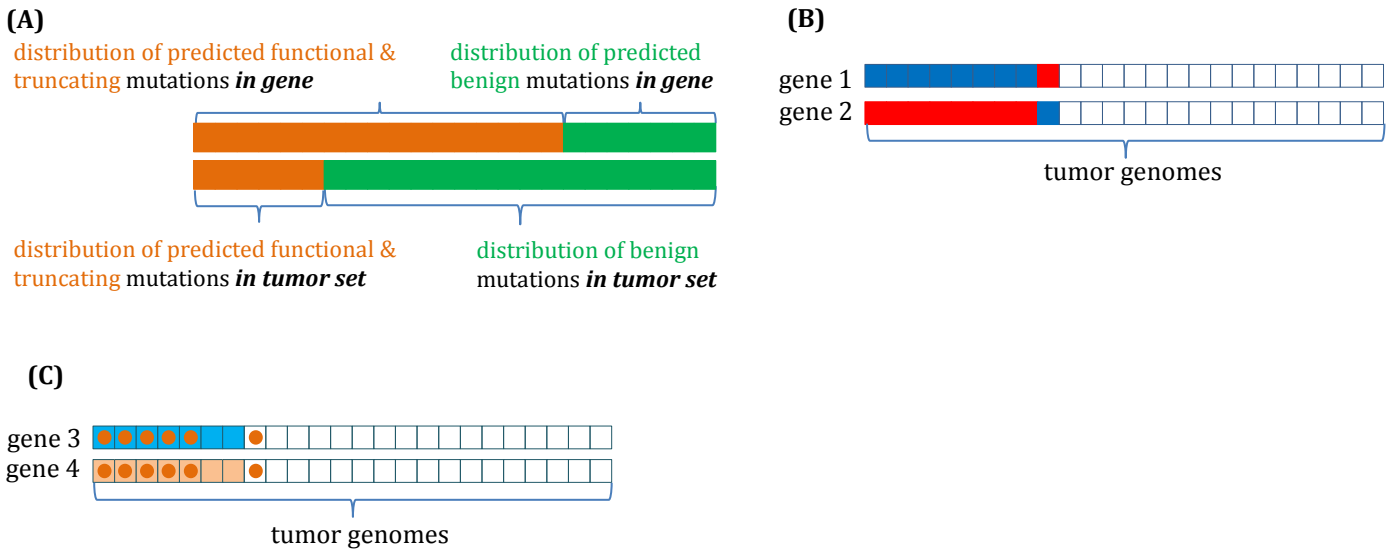
Supplementary Figure 2.



Supplementary Figure 2. Determining signs of positive selection in genomic alterations in cancer. (A) Schematic presentation of distribution of predicted *functional mutations (FM)* [1,2], *truncating mutations (TM) and benign* mutations *(BM) in gene* versus similar distribution in *tumor set*. Enrichment in functional versus benign mutations in a given gene as compared to distribution of functional versus benign mutations across all genes in a tumor set is interpreted as a sign of positive selection [2]. Enrichments in both *FM and FM+TM* mutations are considered. (B) Schematic representation of DNA copy number alterations: *gene 1* represents a gene enriched by *homozygous deletions (DD, blue squares)* versus *amplifications (red squares);* a *gene 2* represents a gene enriched by amplifications (AA) versus *homozygous deletions*. Enrichment in DD vs AA or AA vs DD alterations in a given gene compared to distribution of DD and AA alterations across all genes in a tumor set is interpreted as a sign of positive selection. (C) Schematic presentation of mutation enrichments (FM and TM, brown circles) in regions of DNA copy loss (light blue squares) (gene 3) and in regions of DNA copy gain (brown squares) (gene 4). Enrichment of FM and TM mutations in regions of copy loss/gain versus FM and TM mutations in normal copy regions is interpreted as a sign of positive selection. The statistical significance of enrichment is assessed by Fisher Exact test.

The full list of genes with statistics of genomic alterations and computed signs of selection is presented in Supplement Table S2. Predicted signs of selection were used in analysis of the

carcinogenic role of *myosin IIa* gene (MYH9) . A significant number of genes in this list has not been previously identified as drivers by direct statistical methods as MutSig [3] or MuSiC [4]. For example, MutSig predicted an insignificant Q-value (~0.2) for *myosin IIa* gene (MYH9) that appeared to be a tumor suppressor as we recently reported in *Science* [5]. However, we found a statistically significant enrichment of predicted functional mutations in this gene that is totally consistent with the experimental data and further suggesting the unexpected driver role of MYH9 in HNSCC [5]. Our predictions of driver genes based on pan-cancer TCGA data were also used in nomination of new tumor-suppressors in a recent publication in *Nature Communication* [6].

References

1. Reva, B., Antipin, Y. and Sander, C. (2011) Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research*, **39**, e118.
2. Reva, B. (2013) Revealing selection in cancer using the predicted functional impact of cancer mutations. Application to nomination of cancer drivers. *BMC genomics,* 14 Suppl **3**:S8
3. Getz G, Hofling H, Mesirov JP, Golub TR, Meyerson M, Tibshirani R, Lander ES: Comment on "The consensus coding sequences of human breast and colorectal cancers". Science 2007, 317(5844):1500.
4. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, Mooney TB, Callaway MB, Dooling D, Mardis ER, et al.: MuSiC: identifying mutational significance in cancer genomes. Genome Res 2012, 22(8):1589-1598.
5. Schramek D., Sendoel A., Segal J., Beronja S., Heller E., Oristian D., Reva B., Fuchs E. Direct in Vivo RNAi Screen Unveils Myosin IIa as a Tumor Suppressor of Squamous Cell Carcinomas. Science 17 January 2014: 309-313.
6. Hamilton M, Rajapakshe K, Hartig S, Reva B, McLellan M, Kandoth C, Ding L, Zack T, Gunaratne P, Wheeler D, Coarfa C, McGuire S. Identification of a pan-cancer oncogenic microRNA superfamily anchored by a central core seed motif. Nature Communications 2013 4, Article number: 2730 doi:10.1038/ncomms3730.